case $z_1 < z_0$, since the most interesting case is $z_2 > z_1 \geq 1$ (the amount of the ante).

* The results contained in this paper were obtained in the course of research conducted for The Rand Corporation, Santa Monica, California.

[1] Borel, E., and collaborators, *Traite du Calcul des Probabilités et de ses Applications*, Vol. IV, 2, Chapter V, Paris, 1938.

[2] von Neumann, J., and Morgenstern, O., *The Theory of Games and Economic Behavior*, Chapter 19, Princeton University Press, Princeton, 1947.

---

# *ON FLUCTUATIONS IN COIN-TOSSING**

BY KAI LAI CHUNG AND W. FELLER

CORNELL UNIVERSITY, ITHACA, NEW YORK

Communicated by J. von Neumann, August 4, 1949

In the classical coin-tossing game we have a sequence of independent random variables $X_v$, $v = 1, 2, \ldots$ each taking the values $\pm 1$ with probability 1/2. We are interested in the signs of the partial sums $S_n = \sum_{v=1}^{n} X_v$. To eliminate the zeros of $S_n$ we make the following convention: $S_n$ is "positive" if $S_n > 0$ or if $S_n = 0$ but $S_{n-1} > 0$; otherwise $S_n$ is "negative." The elegance of the results to be announced depends on this convention. Let $N_n$ denote the number of "positive" terms among $S_1$, $\ldots$, $S_n$. We shall confine ourselves to an even $n$, noting only that $0 \leq N_{n+1} - N_n \leq 1$. In the following $r$ and $m$ are positive integers and $P(B|A)$ is the conditional probability of $B$ under the hypothesis $A$.

THEOREM 1.

$$P(N_{2n} = 2r) = \frac{1}{2^{2n}} \binom{2r}{r} \binom{2n-2r}{n-r}.$$

THEOREM 2A.

$$P(N_{2n} = 2r | S_{2n} = 0) = \frac{1}{n+1}.$$

THEOREM 2.

$$P(N_{2n} = 2r | S_{2n} = 2m) = m \binom{2n}{n-m}^{-1} \sum_{i=m}^{r} \binom{2i}{i-m} \binom{2n-2i}{n-i} \frac{1}{i(n-i+1)}.$$

From these theorems it is easy to derive the following *limiting forms*

$$\lim_{n \to \infty} P(N_n \leq \alpha n) = \frac{2}{\pi} \text{arc sin } \alpha^{1/2} \qquad 0 < \alpha < 1 \qquad (a)$$

(with the corresponding density $\pi^{-1}(\alpha - \alpha^2)^{-1/2}$), and

$$\lim_{n \to \infty} P(N_n \leq \alpha n | S_n = \lambda n^{1/2}) = \pi^{1/2} \int_{(1-\alpha)\lambda^2/\alpha} (\lambda^2(\alpha - 1) + \alpha y) y^{-3/2} e^{-y}\, dy \qquad (b)$$

for $0 < \alpha < 1$ and for every fixed $\lambda$ such that $\lambda n^{1/2}$ is a possible value of $S_n$. As $\lambda \to 0$ the last integral tends to $\alpha$ and for $\lambda = 0$ we get in the limit the uniform distribution as suggested by Theorem 2A.

The limiting form (*a*) of Theorem 1 is contained in a result of Erdös and Kac.[1] An analog to Theorem 1 for continuous random variables was proved by E. S. Andersen;[2] however, his result does not hold for arbitrary discrete variables, and Theorem 1 is not contained in Andersen's result.

The most surprising information is contained in Theorem 2A. To see this, note that the distribution of Theorem 1 departs radically from the familiar bell-shaped pattern. The limiting probability density curve resulting from Theorem 1 is U-shaped with the mean at the minimum. More precisely, while $\lim_{u \to \infty} E(N_{2n}/2n) = \frac{1}{2}$ this is the *least* probable value of $N_{2n}/2n$. This means that although the game is a symmetrical one (hence "fair") it is nevertheless more likely that one party "leads" in an overwhelmingly large proportion of the time than that each party leads about half of the time. In Theorem 2A it is plausible that the knowledge that $S_{2n}$ vanishes should reduce the probability of extreme values of $N_{2n}$. It is noteworthy that this is true to the extent of a *uniform* distribution: at a moment when there is a tie all possible guesses about the fraction of time during which one partly has been leading are equally probable. This result contrasts not only with Theorem 1 but even more with a result of Paul Lévy[3] (Corollary 2, pp. 303–304). According to Lévy, if the condition $S_{2n} = 0$ is replaced by the hypothesis that at the $2n$th trial $S_{2n}$ vanishes for the $k$th time, the limiting distribution as $k \to \infty$ of Theorem 2A is again the arc sin law (*a*) and no longer (*b*).

These results should serve as a warning to statisticians who might assume that fluctuation phenomena always follow the bell-shaped pattern and who would easily discover secular trends. If a coin is tossed once a second for a total of 365 days, the probability that one of the players will lead for more than 364 days and 10 hours is about 0.05! However, if it is known that the game concluded at a moment where neither player had a gain or loss, then the probability of such an extended lead is less than 0.0002.

*Proof of Theorem 1.* Let

$$p_{2r}(2n) = P(N_{2n} = 2r) \qquad \text{for } n > 0$$

$$p_0(0) = 1, \qquad p_r(0) = 0 \qquad \text{for } r > 0.$$

$$f_{2k} = P(S_j \neq 0 \text{ for } 1 \leq j < 2k;\ S_{2k} = 0) \qquad \text{for } k \leq 1;\ f_0 = 0$$

$$F(t) = \sum_{k=0}^{\infty} f_{2k}t^{2k} = 1 - (1 - t^2)^{1/2}.$$

Unless $r = 0$ or $r = n$, there is a smallest $k$ such that $1 \leq k < n$ and $S_{2k} = 0$. All $S_j$ with $1 \leq j \leq 2k$, are either all positive or all negative. These considerations lead to the recurrence relation

$$p_{2r}(2n) = \frac{1}{2}\sum_{k=0}^{n} f_{2k}p_{2r-2k}(2n - 2k) + \frac{1}{2}\sum_{k=0}^{n} f_{2k}p_{2r}(2n - 2k). \quad (1)$$

For $r = n$ we must add the probability that $S_{2k} > 0$ for $1 \leq k \leq 2n$, which is equal to

$$\frac{1}{2}\sum_{k=n+1}^{\infty} f_{2k} = \frac{1}{2}\binom{2n}{n}\frac{1}{2^{2n}}. \quad (2)$$

For $r = 0$ we must add the same quantity which now represents the probability that $S_{2k} < 0$ for $k \leq 2n$.

Introducing the generating function

$$g_{2n}(s) = \sum_{r=0}^{2n} p_{2r}(2n)s^{2r},$$

we obtain from formula (1)

$$g_{2n}(s) = \frac{1}{2}\sum_{k=0}^{n} f_{2k}s^{2k}g_{2n-2k}(s) + \frac{1}{2}\sum_{k=0}^{n} f_{2k}g_{2n-2k}(s) + \frac{1 + s^{2n}}{2}\binom{2n}{n}\frac{1}{2^{2n}}. \quad (3)$$

We now introduce the double generating function

$$G_s(t) = \sum_{i=0}^{\infty} g_{2i}(S)t^{2i}.$$

Then from formula (3),

$$G_s(t) = \frac{1}{2}\{G_s(t)F(st) + G_s(t)F(t)\} + \frac{1}{2}\{(1 - t)^{-1/2} + (1 - st)^{-1/2}\}, \quad (4)$$

where $F(t) = 1 - (1 - t^2)^{1/2}$, as implied by formula (2). Solving for $G_s(t)$ we get

$$G_s(t) = \frac{(1 - t)^{-1/2} + (1 - st)^{-1/2}}{(1 - t)^{1/2} + (1 - st)^{1/2}} = \frac{1}{(1 - t)^{1/2}(1 - st)^{1/2}}.$$

The coefficient of $s^{2r}t^{2n}$ is $p_{2r}(2n)$ and is that given in Theorem 1.

*Proof of Theorem 2.* We use the same notations as in the preceding proof except that now

$$p_{2r}(2n) = P(N_{2n} = 2r;\ S_{2n} = 2m).$$

The formula (1) remains valid except that for $r = n$ we must now add

$$W_n = P(S_k > 0 \text{ for } 0 < k \leq n;\ S_{2n} = 2m) = \frac{m}{n}\binom{2n}{n-m}\frac{1}{2^{2n}}$$

by a well-known formula (see reference 4, p. 153). For $r = 0$, formula (1) is valid as it stands with both sides vanishing. Using the generating function

$$W(s) = \sum_{n=m}^{\infty} w_{2n}s^{2n} = \left[\frac{1-(1-s^2)^{1/2}}{s}\right]^m$$

we obtain, instead of formula (4),

$$G_s(t) = {}^1/_2\{G_s(t)F(st) + G_s(t)F(t)\} + W(st).$$

Hence

$$G_s(t) = \frac{2W(st)}{(1-t^2)^{1/2} + (1-s^2t^2)^{1/2}} = 2W(st)\,\frac{(1-s^2t^2)^{1/2} - (1-t^2)^{1/2}}{t^2(1-s^2)}. \tag{5}$$

For $m = 0$, $W(s) \equiv 1$. The coefficient of $s^{2r}t^{2n}(r \leq n)$ in the right-hand side of formula (5) is that of $s^{2r}t^{2n+2}$ in

$$-2(1-t^2)^{1/2}(1-s^2)^{-1}.$$

Thus it is the same for all $r$, $0 \leq r \leq n$. This means that $p_{2r}(2n)$, and consequently also $P(N_{2n} = 2r|S_{2n} = 0)$, is the same for all $r$, $0 \leq r \leq n$. Hence it is equal to $\dfrac{1}{n+1}$. Theorem 2A is proved.

If $m > 0$, the same coefficient is easily seen to be

$$\sum_{i=m}^{r}\left|2\binom{1/2}{n-i+1}\right| w_{2i}.$$

This is equivalent to the formula given in Theorem 2.

* This paper is connected with work under an ONR contract for research in probability theory at Cornell University.

[1] Erdös, P., and Kac, M., "On the Number of Positive Sums of Independent Random Variables," *Bull. Am. Math. Soc.*, **53**, 1011–1020 (1947).

[2] Andersen, E. S., "On the Number of Positive Sums of Random Variables," *Skand. Aktuarietid*, 1950, to appear.

[3] Lévy, P., "Sur certains processus stochastiques homogènes," *Compositio Math.*, **7**, 283–339 (1939).

[4] Uspensky, V., *Introduction to Mathematical Probability*, McGraw-Hill, New York, 1937.